GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:24 ; Search time 210.42 Seconds
(without alignments)
9.750 Million cell updates/sec

Title: US-09-331-631A-21_COPY_32_91
Perfect score: 343
Sequence: 1 TENPCAQRCLQSCQQEPDDL.....DTGATNQRHPPGERTRGRQP 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	343	100.0	614	18	W22149	Peanut allergen Ar
2	343	100.0	614	19	W62834	Arachis hypogaea a
3	318	92.7	626	18	W22150	Peanut allergen Ar
4	318	92.7	626	20	Y15244	Peanut allergen, A
5	318	92.7	626	20	Y25657	Peanut allergen 11
6	93	27.1	605	19	W62838	Glycine max antimi
7	89	25.9	605	20	Y40999	Soybean beta-congl
8	69.5	20.3	625	19	W62830	Macadamia integrif
9	68.5	20.0	666	19	W62829	Macadamia integrif
10	66.5	19.4	477	16	R74175	Human collapsin.
11	66.5	19.4	590	19	W62832	Gossypium hirsutum
12	66.5	19.4	771	16	R71380	Human semaphorin I
13	66.5	19.4	796	19	Y21264	Human semaphorin I
14	64.5	18.8	666	19	W62828	Macadamia integrif
15	64	18.7	1036	21	Y82776	Human chordin rela
16	64	18.7	1036	21	Y53034	Human secreted pro
17	63.5	18.5	1524	20	Y15458	Human laminin gamm
18	60.5	17.6	149	8	P70057	Human insulin rece
19	60.5	17.6	777	20	Y27127	Human brain tissue
20	60.5	17.6	777	21	Y99427	Human PRO1491 (UNQ
21	60.5	17.6	1105	20	Y15459	SEQ ID 5 of WO9919
22	60.5	17.6	1761	20	Y15457	Human laminin beta
23	60	17.5	292	20	W86225	Human VEGF-C trunc
24	60	17.5	297	20	W86224	Human VEGF-C trunc
25	60	17.5	302	20	W86223	Human VEGF-C trunc
26	60	17.5	307	20	W86222	Human VEGF-C trunc
27	60	17.5	318	20	Y08284	Human growth facto
28	60	17.5	350	16	R82686	Vascular endotheli
29	60	17.5	350	20	Y30519	A truncated vascul
30	60	17.5	350	20	Y22321	Truncated human VE
31	60	17.5	399	20	W86237	Human VEGF-C full
32	60	17.5	419	18	W17837	Human foetal liver
33	60	17.5	419	18	W00932	Human Flt4 recepto
34	60	17.5	419	18	W13833	Human vascular end
35	60	17.5	419	18	W11478	Human vascular end
36	60	17.5	419	19	W75751	Vascular endotheli
37	60	17.5	419	19	W75740	Human vascular end
38	60	17.5	419	20	Y30518	Vascular endotheli
39	60	17.5	419	20	Y22320	Full length human
40	60	17.5	419	20	W86203	Human vascular end
41	60	17.5	419	21	Y70982	Human vascular end
42	59.5	17.3	318	20	Y38613	Neisseria gonorrho
43	59.5	17.3	1013	19	W61539	Human cardiac/brai
44	59.5	17.3	1013	19	W40223	Murine mTll protei
45	59.5	17.3	1013	19	W40224	Human tolloid-like

ALIGNMENTS

RESULT 1
W22149
ID W22149 standard; Protein; 614 AA.
XX
AC W22149;
XX
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hI.
XX
KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy;
KW monoclonal antibody; ELISA; analysis; Ara hI.
XX
OS Arachis hypogaea strain Florunner.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= Sig_peptide
FT Protein 23..614
FT /label= Mat_protein
FT Modified-site 521..523
FT /note= "N-glycosylation site"
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US15222.
XX
PR 04-MAR-1996; 96US-0610424.
PR 29-DEC-1995; 95US-0009455.
XX
PA (UYAR-) UNIV ARKANSAS.
XX

PI Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
XX
DR WPI; 1997-363453/33.
DR N-PSDB; T76612.
XX
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in
PT two-site monoclonal antibody based ELISA
XX
PS Claim 31; Page 169; 354pp; English.
XX
CC This polypeptide comprises major peanut allergen Ara hI (W22149).
CC Its sequence was deduced from cDNA clone P17 (T76612), isolated
CC from peanut seed cDNA using a primer (see T76616) based on an
CC isolated Ara hI peptide (see W24206). The sequence shows
CC significant homology with the vicilin family of seed storage
CC proteins of other legumes. The allergen is recognised by serum
CC IgE from a large proportion of individuals with peanut
CC hypersensitivity. Ara hI and Ara hII (see W24164) can be used to
CC raise monoclonal antibodies which are used in a specific two-site
CC MAb ELISA for the detection of Ara hI or Ara hII (claimed). IgE-
CC binding Ara hI antigen epitopes (see W24165-87) may be used in
CC vaccines to protect against allergic reactions to peanut allergens,
CC e.g. anaphylactic shock.
XX
SQ Sequence 614 AA;

Query Match 100.0%; Score 343; DB 18; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.3e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGERTRGRQP 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 32 tenpcaqrclqscqqepddlkqkacesrctkleydprcvydtgatnqrhppgertrgrqp 91

RESULT 2
W62834
ID W62834 standard; Peptide; 614 AA.
XX
AC W62834;
XX
DT 27-OCT-1998 (first entry)
XX
DE Arachis hypogaea antimicrobial protein.
XX
KW antimicrobial protein; infestation; control.
XX
OS Arachis hypogaea.
XX
PN WO9827805-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-AU00874.
XX
PR 20-DEC-1996; 96AU-0004275.
XX
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR WPI; 1998-377279/32.
XX
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
PS Claim 1; Page 55-57; 96pp; English.
XX
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
XX
SQ Sequence 614 AA;

Query Match 100.0%; Score 343; DB 19; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.3e-32;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGERTRGRQP 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 32 tenpcaqrclqscqqepddlkqkacesrctkleydprcvydtgatnqrhppgertrgrqp 91

RESULT 3
W22150
ID W22150 standard; Protein; 626 AA.
XX
AC W22150;
XX
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hI.
XX
KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy;
KW monoclonal antibody; ELISA; analysis; Ara hI.
XX
OS Arachis hypogaea strain Florunner.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= Sig_peptide
FT Protein 23..626
FT /label= Mat_protein
FT Modified-site 521..523
FT /note= "N-glycosylation site"
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US15222.
XX
PR 04-MAR-1996; 96US-0610424.
PR 29-DEC-1995; 95US-0009455.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
XX
DR WPI; 1997-363453/33.
DR N-PSDB; T76613.
XX
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in
PT two-site monoclonal antibody based ELISA
XX
PS Claim 31; Page 172; 354pp; English.
XX
CC This polypeptide comprises major peanut allergen Ara hI (W22149).
CC Its sequence was deduced from cDNA clone P41b (T76613), isolated
CC from peanut seed cDNA using a primer (see T76616) based on an
CC isolated Ara hI peptide (see W24206). The sequence shows
CC significant homology with the vicilin family of seed storage
CC proteins of other legumes. The allergen is recognised by serum
CC IgE from a large proportion of individuals with peanut
CC hypersensitivity. Ara hI and Ara hII (see W24164) can be used to
CC raise monoclonal antibodies which are used in a specific two-site
CC MAb ELISA for the detection of Ara hI or Ara hII (claimed). IgE-
CC binding Ara hI antigen epitopes (see W24165-87) may be used in
CC vaccines to protect against allergic reactions to peanut allergens,
CC e.g. anaphylactic shock.
XX
SQ Sequence 626 AA;

```
  Query Match             92.7%;  Score 318;  DB 18;  Length 626;
  Best Local Similarity   90.6%;  Pred. No. 1.2e-29;
  Matches   58;  Conservative    0;  Mismatches    2;  Indels    4;  Gaps    1;

Qy       1 TENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYD----TGATNQRHPPGERTR 56
           ||||||||||||||||||||||||||||||||||||||||||    || |||| |||||||
Db      34 tenpcaqrclqscqqepddlkqkacesrctkleydprcvydprghtgttnqrsppgertr 93

Qy      57 GRQP 60
           ||||
Db      94 grqp 97


RESULT   4
Y15244
ID   Y15244 standard; Protein; 626 AA.
XX
AC   Y15244;
XX
DT   09-NOV-1999  (first entry)
XX
DE   Peanut allergen, Ara h 1, amino acid sequence.
XX
KW   allergy; immune response; transgenic; allergen; epitope;
KW   immunoglobulin E; Ig E; binding site; peanut.
XX
OS   Arachis hypogea.
XX
PN   WO9938978-A1.
XX
PD   05-AUG-1999.
XX
PF   29-JAN-1999;   99WO-US02031.
XX
PR   27-AUG-1998;   98US-0141220.
PR   31-JAN-1998;   98US-0073283.
PR   13-FEB-1998;   98US-0074590.
PR   13-FEB-1998;   98US-0074624.
PR   13-FEB-1998;   98US-0074633.
XX
PA   (SOSI/) SOSIN H.
PA   (UYAR-) UNIV ARKANSAS.
PA   (UYNY ) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX
PI   Bannon GA,  Burks AW,  Sampson HA,  Sosin H;
XX
DR   WPI; 1999-479189/40.
DR   N-PSDB; Z06382.
XX
PT   Modified allergen with reduced IgE binding, useful for treating e.g.
PT   allergies
XX
PS   Disclosure; Page 35-37; 46pp; English.
XX
CC   This is the amino acid sequence of the Ara h 1 protein from Arachis
CC   hypogea. The Ara h 1 protein has 23 IgE (Immunoglobulin E) binding
CC   epitopes, four of which are immunodominant (Y15247, Y15249, Y15250 and
CC   Y15263).
CC   By modifying the IgE binding sites the ability of the allergen to
CC   provoke an immune response is downregulated. The epitopes of the IgE
CC   binding sites can therefore be modified in genetically engineered plants
CC   and animals to elicit less of an allergic response.
XX
SQ   Sequence   626 AA;


  Query Match             92.7%;  Score 318;  DB 20;  Length 626;
  Best Local Similarity   90.6%;  Pred. No. 1.2e-29;
  Matches   58;  Conservative    0;  Mismatches    2;  Indels    4;  Gaps    1;

Qy       1 TENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYD----TGATNQRHPPGERTR 56
           ||||||||||||||||||||||||||||||||||||||||||    || |||| |||||||
Db      34 tenpcaqrclqscqqepddlkqkacesrctkleydprcvydprghtgttnqrsppgertr 93

Qy      57 GRQP 60
           ||||
Db      94 grqp 97


RESULT   5
Y25657
ID   Y25657 standard; protein; 626 AA.
XX
AC   Y25657;
XX
DT   30-SEP-1999  (first entry)
XX
DE   Peanut allergen 1168391 Ara h I protein fragment.
XX
KW   Major histocompatibility complex; class II; desensitising; human;
KW   allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW   chiromidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW   screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW   cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW   mice; gerbil; vaccine; treatment; prevention; hypersensitivity; peanut.
XX
OS   Arachis hypogea.
XX
PN   WO9934826-A1.
XX
PD   15-JUL-1999.
XX
PF   11-JAN-1999;   99WO-GB00080.
XX
PR   21-SEP-1998;   98GB-0020474.
PR   09-JAN-1998;   98GB-0000445.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Kay AB,  Larche M;
XX
DR   WPI; 1999-458255/38.
XX
PT   Desensitizing patients to polypeptide allergens
XX
PS   Example 6; Page 70-71; 117pp; English.
XX
CC   This invention describes a novel method of desensitizing a patient to a
CC   polypeptide allergen and comprises administering to the patient a peptide
CC   derived from the allergen where restriction to a MHC Class II molecule
CC   possessed by the patient can be demonstrated for the peptide and the
CC   peptide is able to induce a late phase response in an individual who
CC   possesses the MHC Class II molecule. The methods can be used for
CC   desensitising patients to allergens present in e.g. grass, tree and weed
CC   (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC   the chiromidae (non-biting midges), spiders and mites, housefly, fruit
CC   fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC   non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC   Tenibrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC   sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC   produce immunological vaccines which may be used to prevent and/or treat
CC   conditions involving hypersensitivity to allergens. This sequence
CC   represents a peanut (Arachis hypogea) allergen 1168391 Ara h I.
XX
SQ   Sequence   626 AA;


  Query Match             92.7%;  Score 318;  DB 20;  Length 626;
  Best Local Similarity   90.6%;  Pred. No. 1.2e-29;
  Matches   58;  Conservative    0;  Mismatches    2;  Indels    4;  Gaps    1;

Qy       1 TENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYD----TGATNQRHPPGERTR 56
           ||||||||||||||||||||||||||||||||||||||||||    || |||| |||||||
Db      34 tenpcaqrclqscqqepddlkqkacesrctkleydprcvydprghtgttnqrsppgertr 93
```

```
Qy      57 GRQP 60
           ||||
Db      94 grqp 97


RESULT   6
W62838
ID   W62838 standard; Protein; 605 AA.
XX
AC   W62838;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Glycine max antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Glycine max.
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 63-65; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   605 AA;


  Query Match             27.1%;  Score 93;  DB 19;  Length 605;
  Best Local Similarity   31.4%;  Pred. No. 0.0035;
  Matches   22;  Conservative    8;  Mismatches   16;  Indels   24;  Gaps    3;

Qy       2 ENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYD-----------PRCVYDTGATNQRHP 50
           ||   :|||||  | |  : :|| :||  |: :           ||          :||
Db      31 enpkhnkclqscnserdsyrnqacharcnllkvekeeceegeiprpr-------prpqhp 83

Qy      51 ------PGER 54
                |||:
Db      84 erepqqpgek 93


RESULT   7
Y40999
ID   Y40999 standard; protein; 605 AA.
XX
AC   Y40999;
XX
DT   06-DEC-1999  (first entry)
XX
DE   Soybean beta-conglycinin protein sequence.
XX
KW   Peanut; allergen; Ara H 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW   allergic reaction; soybean; beta-conglycinin.
XX
OS   Glycine max.
XX
PN   WO9945961-A1.
XX
PD   16-SEP-1999.
XX
PF   12-MAR-1999;   99WO-US05494.
XX
PR   12-MAR-1998;   98US-0077763.
PR   11-MAR-1999;   99US-0077763.
XX
PA   (UYAR-) UNIV ARKANSAS.
XX
PI   Burks W,  Helm RM,  Cockrell G,  Bannon GA,  Stanley JS,  Shin DS;
PI   Sampson H,  Compadre CM,  Huang SK,  Maleki SJ,  Kopper RA;
XX
DR   WPI; 1999-551218/46.
XX
PT   Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT   animal from allergic reaction -
XX
PS   Disclosure; Fig 33A-B; 193pp; English.
XX
CC   The invention provides a tertiary structure for the peanut allergen
CC   Ara H 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC   epitopes. The invention also provides an isolated recombinant peanut
CC   allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC   allergen Ara h 3. Molecules of the invention are used to protect a host
CC   animal from allergic reaction, particularly using a modified allergen
CC   which is less reactive with IgE. The invention may also be used to
CC   ensure that the allergen is not introduced into genetically modified
CC   food. The present sequence represents a soybean beta-conglycinin protein
CC   sequence.
XX
SQ   Sequence   605 AA;


  Query Match             25.9%;  Score 89;  DB 20;  Length 605;
  Best Local Similarity   30.0%;  Pred. No. 0.01;
  Matches   21;  Conservative    9;  Mismatches   16;  Indels   24;  Gaps    3;

Qy       2 ENPCAQRCLQSCQQEPDDLKQKACESRCTKLEYD-----------PRCVYDTGATNQRHP 50
           :||   :|||||  | |  : :|| :||  |: :           ||          :||
Db      31 knpkhnkclqscnserdsyrnqacharcnllkvekeeckkgeiprpr-------prpqhp 83

Qy      51 ------PGER 54
                |||:
Db      84 erepqqpgek 93


RESULT   8
W62830
ID   W62830 standard; Protein; 625 AA.
XX
AC   W62830;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Macadamia integrifolia antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Macadamia integrifolia.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..28
FT                   /note= "signal peptide"
FT   Protein         29..666
FT                   /note= "mature protein"
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
```

```
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
DR   N-PSDB; V42316.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 43-45; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   625 AA;


  Query Match             20.3%;  Score 69.5;  DB 19;  Length 625;
  Best Local Similarity   30.0%;  Pred. No. 2.1;
  Matches   12;  Conservative   10;  Mismatches   11;  Indels    7;  Gaps    1;

Qy      5 CAQRCLQSCQQEPDDLKQK-------ACESRCTKLEYDPR 37
          | :|| : |::| :  :|:        |: || : | :||
Db     61 cqrrckeiceeeeeynrqrdpqqqyeqcqkrcqrretepr 100


RESULT   9
W62829
ID   W62829 standard; Protein; 666 AA.
XX
AC   W62829;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Macadamia integrifolia antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Macadamia integrifolia.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..28
FT                   /note= "signal peptide"
FT   Protein         29..666
FT                   /note= "mature protein"
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
DR   N-PSDB; V42311.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 39-41; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   666 AA;


  Query Match             20.0%;  Score 68.5;  DB 19;  Length 666;
  Best Local Similarity   30.0%;  Pred. No. 3;
  Matches   12;  Conservative   10;  Mismatches   11;  Indels    7;  Gaps    1;

Qy      5 CAQRCLQSCQQEPDDLKQK-------ACESRCTKLEYDPR 37
          | :|| : |::| :  :|:        |: || : | :||
Db    102 cqrrckeiceeeeeynrqrdpqqqyeqcqercqrhetepr 141


RESULT  10
R74175
ID   R74175 standard; Protein; 477 AA.
XX
AC   R74175;
XX
DT   01-NOV-1995  (first entry)
XX
DE   Human collapsin.
XX
KW   Collapsin; antibody; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Binding-site    9..19
FT                   /note= "antibody binding site"
FT   Binding-site    51..65
FT                   /note= "antibody binding site"
XX
PN   US5416197-A.
XX
PD   16-MAY-1995.
XX
PF   15-OCT-1993;   93US-0136922.
XX
PR   15-OCT-1993;   93US-0136922.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Luo Y,  Raper JA;
XX
DR   WPI; 1995-193478/25.
DR   N-PSDB; Q92331.
XX
PT   New antibody to human collapsin - used to inhibit the activity of
PT   collapsin, to induce neurite out-growth and to treat individuals with
PT   nerve damage.
XX
PS   Claim 2; Columns 15-18; 11pp; English.
XX
CC   An antibody capable of specifically binding at least a portion of
CC   the collapsin protein can be used to purify human collapsin and
CC   to inhibit the activity of the protein.  It can be used to induce
CC   neurite outgrowth by neuronal cells and to treat individuals
CC   suffering from nerve damage.
XX
SQ   Sequence   477 AA;


  Query Match             19.4%;  Score 66.5;  DB 16;  Length 477;
  Best Local Similarity   35.2%;  Pred. No. 3.6;
  Matches   19;  Conservative    5;  Mismatches   23;  Indels    7;  Gaps    2;

Qy      6 AQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGERTRGRQ 59
          ||  |  |     |:  |||   |  |  || | :|  | ::  |  :|   ||
Db    278 aqlplhrc-----diygkacaecc--lardpycawdgsacsryfptakrrtrrq 324
```

```
RESULT 11
W62832
ID   W62832 standard; Protein; 590 AA.
XX
AC   W62832;
XX
DT   27-OCT-1998 (first entry)
XX
DE   Gossypium hirsutum antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Gossypium hirsutum.
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR   WPI; 1998-377279/32.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 49-51; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   590 AA;


  Query Match             19.4%;  Score 66.5;  DB 19;  Length 590;
  Best Local Similarity   31.6%;  Pred. No. 4.5;
  Matches   18;  Conservative   10;  Mismatches   22;  Indels    7;  Gaps    3;

Qy       5 CAQRCLQSCQQEPDDLKQ--KACESRCTKLEYDP----RCVYDTGATNQRHP-PGER 54
           | ||||:  :||    ::  : |:  | : |  |    :|| :     | :|  |||
Db     105 cqgrclkrfeqeqqqsqrqfqecqqhchqqeqrpekkqqcvrecrekyqenpwrger 161


RESULT 12
R71380
ID   R71380 standard; Protein; 771 AA.
XX
AC   R71380;
XX
DT   21-NOV-1995 (first entry)
XX
DE   Human semaphorin III protein.
XX
KW   Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW   variola major virus; smallpox; semaphorin receptor binding activity;
KW   modulation; nerve cell growth; immune response; viral pathogenesis;
KW   neurological disease; neuro-regeneration; oncological infection.
XX
OS   Homo sapiens.
XX
PN   WO9507706-A.
XX
PD   23-MAR-1995.
XX
PF   13-SEP-1994;   94WO-US10151.
XX
PR   13-SEP-1993;   93US-0121713.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI   O'Connor T;
XX
DR   WPI; 1995-131177/17.
DR   N-PSDB; Q87442.
XX
PT   New class of semaphorin peptide(s) and polypeptide(s) - are
PT   potent modulators of nerve cell growth and regeneration
XX
PS   Example 2; Page 60-63; 101pp; English.
XX
CC   The sequence of the human semaphorin III protein. The proteins
CC   encoded by the grasshopper semaphorin I (Q87441), human semaphorin III,
CC   vaccinia virus semaphorin IV (Q87443), Drosophila semaphorin I and II
CC   (Q87444-5), Tribolium semaphorin I (Q87446) or variola major (smallpox)
CC   virus semaphorin IV (Q87447) genes were used to generate a series of
CC   peptides (R70370-R70418), which retain semaphorin receptor binding
CC   activity. The semaphorin derived or semaphorin receptor derived peptides
CC   are potent modulators of nerve cell growth, immune responsiveness and
CC   viral pathogenesis. They can be used in diagnosis and treatment of
CC   neurological disease and neuro-regeneration, immune modulation and
CC   diagnosis and treatment of viral and oncological infection and diseases.
XX
SQ   Sequence   771 AA;


  Query Match             19.4%;  Score 66.5;  DB 16;  Length 771;
  Best Local Similarity   35.2%;  Pred. No. 6;
  Matches   19;  Conservative    5;  Mismatches   23;  Indels    7;  Gaps    2;

Qy       6 AQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGERTRGRQ 59
           ||  |  |     |:  |||   |  |  || | :|  | ::  |  :|   ||
Db     510 aqlplhrc-----diygkacaecc--lardpycawdgsacsryfptakrrtrrq 556


RESULT 13
Y21264
ID   Y21264 standard; Protein; 796 AA.
XX
AC   Y21264;
XX
DT   22-JUL-1999 (first entry)
XX
DE   Human semaphorin III wild type protein fragment 1.
XX
KW   Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW   frameshift mutation; age-related disease; neurodegenerative disorder;
KW   Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW   Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW   diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW   ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW   neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW   glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW   bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW   high mobility group protein-C; neuroendocrine specific protein A.
XX
OS   Homo sapiens.
XX
PN   WO9845322-A2.
XX
PD   15-OCT-1998.
XX
PF   02-APR-1998;   98WO-IB00705.
XX
PR   10-APR-1997;   97US-0043163.
XX
PA   (UYUT-) RIJKSUNIV UTRECHT.
PA   (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA   (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
```

```
PI   Burbach JPH,  Grosveld FG,  Van Leeuwen FW;
XX
DR   WPI; 1998-609901/51.
DR   N-PSDB; X75767.
XX
PT   Diagnosing disease by detecting frameshift mutations in RNA or
PT   corresponding protein mutations - used to diagnose cancer and
PT   neurological diseases, particularly Alzheimer's disease, and also
PT   for treatment and prevention with specific ribozymes or wild-type
PT   RNA
XX
PS   Disclosure; Figure 16; 258pp; English.
XX
CC   This invention describes a novel method for the diagnosis of a disease
CC   caused by, or associated with, an RNA molecule that has a frameshift
CC   mutation. The method is used to diagnose age-related diseases, especially
CC   cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC   disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC   multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC   and many others listed) or susceptibility to these disorders. The method
CC   allows a definitive diagnosis of Alzheimer's disease in living patients,
CC   at an early stage. It is based on the observation that disease may be
CC   caused by mutations in RNA rather than DNA. The invention describes the
CC   used of neuronal system RNA molecules, specifically proteins including
CC   beta-amyloid precursor protein (beta-APP), the microtubule associated
CC   proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC   associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC   neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
CC   protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC   2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC   protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
SQ   Sequence   796 AA;


  Query Match             19.4%;  Score 66.5;  DB 19;  Length 796;
  Best Local Similarity   35.2%;  Pred. No. 6.2;
  Matches   19;  Conservative    5;  Mismatches   23;  Indels    7;  Gaps    2;

Qy       6 AQRCLQSCQQEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGERTRGRQ 59
           ||  |  |     |:  |||   |  |  || | :|  | ::  |  :|   ||
Db     535 aqlplhrc-----diygkacaecc--lardpycawdgsacsryfptakrrtrrq 581


RESULT  14
W62828
ID   W62828 standard; Protein; 666 AA.
XX
AC   W62828;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Macadamia integrifolia antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Macadamia integrifolia.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..28
FT                   /note= "signal peptide"
FT   Protein         29..666
FT                   /note= "mature protein"
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
DR   N-PSDB; V42310.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 34-36; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   666 AA;


  Query Match             18.8%;  Score 64.5;  DB 19;  Length 666;
  Best Local Similarity   27.5%;  Pred. No. 8.8;
  Matches   11;  Conservative   10;  Mismatches   12;  Indels    7;  Gaps    1;

Qy       5 CAQRCLQSCQQEPDDLKQK-------ACESRCTKLEYDPR 37
           | :|| : |::| :  :|:         |:  | : | :||
Db     102 cqrrckeiceeeeeynrqrdpqqqyeqcqkhcqrretepr 141


RESULT  15
Y82776
ID   Y82776 standard; Protein; 1036 AA.
XX
AC   Y82776;
XX
DT   19-JUN-2000  (first entry)
XX
DE   Human chordin related protein (Clone dj167_19).
XX
KW   Chordin related protein; cartilage; bone; connective tissue;
KW   periodontal disease; osteoporosis; burn; incision; ulcer; neuron;
KW   astrocyte; glial cell; transplantation; nerve; epidermis; muscle;
KW   liver; brain; lung; cardiac; pancreas; kidney; growth;
KW   differentiation; TGF-Beta; angiogenesis; chemotaxis;
KW   chemoattraction; collagen synthesis; fibrosis; cell adhesion;
KW   cell migration; fertility; reproduction; haematopoiesis;
KW   erythroid cell; tumour; dietary supplement; growth medium.
XX
OS   Homo sapiens.
XX
PN   WO200009551-A1.
XX
PD   24-FEB-2000.
XX
PF   10-AUG-1999;   99WO-US18117.
XX
PR   10-AUG-1998;   98US-0095880.
PR   06-MAY-1999;   99US-0306111.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Jacobs K,  Mccoy JM,  Lavallie ER,  Collins-racie LA,  Merberg D;
PI   Treacy M,  Diblasio-smith E,  Widom A;
XX
DR   WPI; 2000-205978/18.
DR   N-PSDB; Z93172.
XX
PT   New polynucleotides encoding secreted human proteins, useful for
PT   treating e.g. broken bones, craniofacial defects, peridontal disease,
PT   osteoporosis, burns, incisions or ulcers  -
XX
PS   Claim 21; Page 94-98; 105pp; English.
XX
CC   The human chordin related protein and polynucleotides encoding them
CC   are predicted to have biological activities which would make them
```

CC suitable for treating, preventing or ameliorating medical conditions
CC which involve defects in cartilage, bone or connective tissue
CC formation and damage to cartilage, bone or connective tissue, e.g.
CC broken bones, congenital, trauma-induced, or
CC oncologic-resection-induced craniofacial defects, periodontal
CC disease, defects in the periodontal ligament or attachment apparatus,
CC damage to the periodontal ligament or attachment apparatus,
CC osteoporosis, burns, incisions or ulcers. The proteins may also
CC affect neuronal, astrocytic, and glial cell survival and therefore be
CC useful in transplantation and treatment of conditions exhibiting a
CC decrease in neuronal survival and repair. The proteins may also be
CC useful for the treatment of conditions related to other types of
CC tissue, such as nerve, epidermis, muscle, and other organs such as
CC liver, brain, lung, cardiac, pancreas, and kidney tissue. The
CC proteins may further be useful for the treatment of relatively
CC undifferentiated cell populations, such as embryonic cells, or stem
CC cells, to enhance growth and/or differentiation of the cells.
CC The proteins may also have other useful properties characteristic of
CC the TGF-beta superfamily of proteins. Such properties include
CC angiogenic, chemotactic, and/or chemoattractant properties, and
CC effects on cells including induction or inhibition of collagen
CC synthesis, fibrosis, differentiation responses, cell proliferative
CC responses, and responses involving cell adhesion, migration, and
CC extracellular matrices. These properties make the proteins potential
CC agents for wound healing, reduction of fibrosis, and reduction of
CC scar tissue formation. Chordin-related proteins may also be useful
CC for advancement of the onset of fertility in sexually immature
CC mammals, so as to increase the lifetime reproductive performance of
CC domestic animals such as cows, sheep and pigs. Chordin-related
CC proteins may also be useful in modulating hematopoiesis by inducing
CC the differentiation of erythroid cells, for suppressing the
CC development of gonadal tumors, or for augmenting the activity of
CC BMPs. The proteins may also have value as a dietary supplement, or
CC as a component of cell culture media.
XX
SQ Sequence 1036 AA;

Query Match 18.7%; Score 64; DB 21; Length 1036;
Best Local Similarity 29.1%; Pred. No. 16;
Matches 16; Conservative 8; Mismatches 23; Indels 8; Gaps 2;

```
Qy       2 ENPCAQRCLQS---CQQEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGE 53
           | |    || :    ::|  |  :  ||     ::: |||  |:      ||||
Db     153 efpsqdmclsalkrieeekpdcskarce-----vqfsprcpedsvliegyappge 202
```

Search completed: March 1, 2001, 15:47:25
Job time: 250 sec